

ABSTRACT

The present invention provides systems and methods for obtaining information from a networked system utilizing a distributed web crawler. The distributed nature of clients of a server is leveraged to provide fast and accurate web crawling data.

Information gathered by a server's web crawler is compared to data retrieved by clients of the server to update the crawler's data. In one instance of the present invention, data comparison is achieved by utilizing information disseminated *via* a search engine results page. In another instance of the present invention, data validation is accomplished by client dictionaries, emanating from a server, that summarize web crawler data. The present invention also facilitates data analysis by providing a means to resist spoofing of a web crawler to increase data accuracy.